Analyse de **données**

Mathieu Laubignat
CISA, Auditeur informatique

« Les techniques d'analyse des données, utilisées depuis 1998 au sein de l'Audit Informatique du Groupe La Poste, constituent un outil d'audit puissant et indispensable pour la réalisation des missions. Cet article présente la démarche utilisée dans ce cadre, et fournit quelques exemples concrets de mise en œuvre. »

1. Présentation succincte de l'Audit Informatique du Groupe La Poste

L'audit Informatique du Groupe La Poste est une structure dont le rôle consiste à évaluer les processus de management des risques et de gouvernance des systèmes d'information du Groupe et à apporter des conseils pour renforcer leur efficacité. Le périmètre du service s'étend à la maison mère et à l'ensemble des filiales du Groupe. Les missions d'audit réalisées peuvent être de différentes natures : audit de SI, audit de sécurité, audit de thématique... Les techniques d'analyse de données, utilisées depuis 1998 au sein de l'Audit Informatique du Groupe La Poste, constituent un outil d'audit puissant dans la panoplie des auditeurs pour la réalisation de ces missions.

2. Pourquoi l'analyse de données ?

L'analyse de données, au sens large, consiste à récolter et exploiter des données quelles que soient leur origine (documents papier, chiffres clés, relevages ou comptages manuels, système d'information...) pour étayer une opinion d'audit. Dans le cas qui nous intéresse, il s'agit plus particulièrement de l'exploitation des données du système d'information à des fins d'audit, on parle alors d'« audit par les données », pratiqué au moyen de techniques d'audit assisté par ordinateur (TAAO ou CAATs).

Les nouvelles normes pour la pratique professionnelle de l'audit interne applicables depuis le 1er janvier 2004 précisent que les auditeurs internes doivent posséder une bonne connaissance des principaux risques et contrôles liés aux technologies de l'information et des techniques d'audit informatisées susceptibles d'être mises en œuvre dans le cadre des travaux qui leur sont confiés (norme 1210.A3). De même, la norme CNCC 2-302, promue par la CNCC (Compagnie Nationale des Commissaires aux Comptes) en 2003, consacre la nécessité d'évaluer systématiquement les risques liés au système d'information et recommande l'utilisation des techniques d'audit assistées par ordinateur. Enfin, la Loi de Sécurité Financière (LSF) du 1er août 2003 induit de nouvelles obligations de contrôle interne pour les entreprises qui doivent rendre compte dans un rapport annuel des procédures de contrôle interne mise en place pour garantir une meilleure transparence. Cette obligation couvre le contrôle interne informatique en tant que rouage essentiel du contrôle interne général.

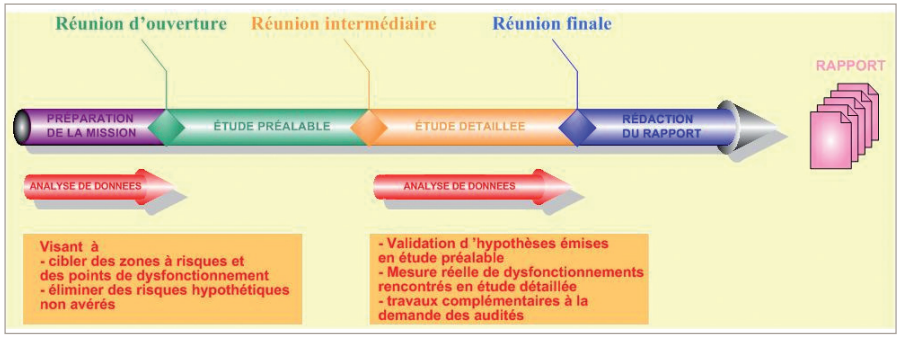
Pour l'auditeur, l'analyse de données accroît considérablement la valeur ajoutée de ses travaux et, de fait, l'image de marque de l'audit. En effet, elle permet de gagner en productivité, d'étendre le périmètre des tests et de fiabiliser l'audit. Il n'y a plus besoin de contrôles manuels longs et fastidieux, les tests portent sur l'exhaustivité de la population, laissant de côté le sondage et l'extrapolation. Enfin, les travaux d'analyse de données permettent d'obtenir des éléments probants opposables avec un chiffrage direct des anomalies à travers une démarche auditable.

3. L'analyse de données, mode d'emploi !

Deux approches peuvent être envisagées pour positionner les travaux d'analyse de données dans le processus d'audit (Fig. 1) :

- En amont de la mission pour cibler les zones à risques ou éliminer des risques hypothétiques non avérés.
- Dans la phase d'investigation pour apporter des constats chiffrés ou étayer une opinion d'audit.

En termes d'organisation, différentes approches sont également envisageables. Le service peut faire le choix de la généralisation ou celui de la spécialisation. Il est vrai que la première option demande un investissement plus important en termes de formation et d'acquisition de logiciels mais la


Réunion d'ouverture
Réunion intermédiaire
Réunion finale
RAPPORT
PREPARATION DE LA MISSION
ÉTUDE PRÉALABLE
ÉTUDE DETAILLEE
REDACTION DU RAPPORT
ANALYSE DE DONNEES
ANALYSE DE DONNEES
Visant à
- cibler des zones à risques et des points de dysfonctionnement
- éliminer des risques hypothétiques non avérés
- Validation d 'hypothèses émises en étude préalable
- Mesure réelle de dysfonctionnements rencontrés en étude détaillée
- travaux complémentaires à la demande des audités

> Fig. 1 : les travaux d'analyse de données dans le processus d'audit

deuxième qui paraît plus économique présente cependant certains désavantages. En effet, un auditeur cantonné aux travaux d'analyse de données risque de se lasser rapidement. D'autre part, si les compétences en la matière sont concentrées sur un nombre restreint d'individus, il y a un risque plus important en cas d'absence ou de départ.

C'est pourquoi, l'Audit Informatique du Groupe La Poste a fait le choix de développer cette compétence sur l'ensemble des auditeurs du pôle ASI. Ainsi, les nouveaux arrivants doivent maîtriser les techniques d'analyses de données au même titre que les travaux d'audit classiques. Cette montée en compétence s'effectue à travers une formation de base assurée par les auditeurs plus expérimentés en interne. De plus, le service dispose d'un outil destiné à capitaliser l'expérience engrangée en matière d'analyse de données. Au sein du Groupe, des services d'audit « généralistes » ont mis en œuvre la même démarche.

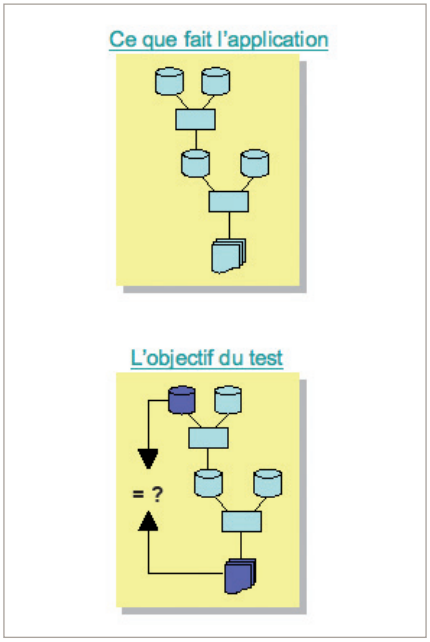
Les travaux d'analyse de données doivent être structurés. En conséquence, le processus se décompose en plusieurs phases :

Phase 1 : Prise de connaissance

Cette phase constitue le préalable à l'ensemble des travaux d'audit. Elle consiste à identifier les acteurs du SI (MOA, MOE,...), connaître le domaine et les objectifs attendus du SI et obtenir sa description (Cartographie applicative, chronologie des traitements, modèle de données,...).

Phase 2 : Définition des scénarios d'analyse

À partir des objectifs de la mission, les tests les plus pertinents sont identifiés. Il peut s'agir de vérifier l'application d'une procédure ou l'implémentation d'une règle de gestion, de vérifier une piste d'audit...



> Fig. 2 : tests d'analyse de données – comparaison des entrées et sorties

Pour valider les scénarios d'analyse, il faut tenir compte des capacités de l'outil d'analyse utilisé et réaliser une approche de faisabilité en considérant la complexité, la volumétrie et la disponibilité des données (planning compatible, accord de l'audité).

Phase 3 : Récupération des données

Après validation et en fonction des objectifs des tests, les données utiles aux travaux d'analyse sont identifiées, les fichiers de départ qui contiendront les champs nécessaires sont définis et arrêtés à une date précise. Il convient alors de transmettre à l'audité par écrit les objectifs de test et les données choisies. Ainsi, celui-ci est informé du périmètre d'extraction et peut valider la sélection, la date, le mode de transmission, le format et la structure des fichiers.

Pour l'extraction, différentes approches sont envisageables en fonction de l'interlocuteur (propriétaire des données, maîtrise d'œuvre, exploitant), du mode employé (mieux vaut privilégier une extraction des données de production à l'état brut mais il peut s'agir également du résultat d'une requête ou de données obtenues à partir d'un infocentre).

Phase 4 : Préparation des données

Une fois les données récupérées, il faut organiser l'environnement de travail (espace d'importation / espace de travail / espace de résultats) et y déposer les fichiers en prenant soin de faire une archive des données de départ afin de pouvoir repartir de zéro le cas échéant.

Ensuite, l'importation des fichiers est réalisée sans oublier d'effectuer un contrôle immédiat pour chaque fichier (validation des totaux de contrôle et du nombre d'enregistrements, contrôles de cohérence,...) dans l'outil d'analyse. Il peut s'agir d'outils de type base de données (Access, Visual Foxpro,....), de logiciels « généraux » (IDEA, ACL), ou encore d'un tableur de type Excel. Cependant, les logiciels « généraux » allient une capacité de traitement importante à une mise en œuvre rapide tout en garantissant l'intégrité des données et la conservation de la piste d'audit (Fig. 3). À titre indicatif, l'Audit Informatique du Groupe La Poste utilise IDEA depuis 1998. Il reste maintenant à formaliser les programmes de travail.

Caractéristiques	Tableurs	Bases de données	ACL/IDEA
Capacité volumétrique	Limitée (65536 pour Excel)	Illimitée	Illimitée
Capacité de traitement	Limitée	Illimitée	Suffisante
Types des fichiers importés	Majorité des fichiers ASCII	Majorité des fichiers ASCII	Tout type de fichiers
Intégrité des données	Impossible	À réaliser	Garantie
Piste d'audit	Impossible	À réaliser	Garantie

> Fig. 3 : Comparaison des principaux outils d'analyse de données

Attention, l'étape de préparation des données peut être la plus fastidieuse si les étapes de prise de connaissance et de définition des scénarios n'ont pas été correctement réalisées. En effet, une mauvaise définition des objectifs et des données cibles peut aboutir à des demandes d'extraction complémentaires ou des travaux de remise en forme des fichiers obtenus qui allongent significativement le processus.

Phase 5 : Réalisation des tests

Pour chaque test, les étapes mises en œuvre sont décrites et notamment les fichiers ou champs créés. Il faut également prévoir des modalités de contrôle pour vérifier les résultats obtenus.

La réalisation d'un document de suivi chronologique des travaux (description de ce qui est fait et ce qui reste à faire) constitue également un élément important durant cette phase. Cette démarche peut permettre éventuellement de prioriser certains travaux en recadrant par la suite avec l'audité.

Phase 6 : Interprétation et présentation des résultats

La dernière phase consiste à analyser et valider les résultats par rapport à ceux attendus en relation avec les acteurs du SI audité. Puis viennent la conclusion et la restitution des résultats sous forme graphique de préférence (attention, il est parfois difficile de représenter certains résultats d'analyse de données). Pour gagner en valeur ajoutée dans les conclusions et les recommandations émises, il est intéressant pour l'équipe d'audit de collaborer autant que possible avec des spécialistes du métier audité.

Si l'analyse n'aboutit pas, il faut en informer les acteurs, identifier la cause et éventuellement prévoir une solution de contournement. Enfin, une synthèse des travaux a *posteriori* permet d'identifier les bonnes pratiques à généraliser lors des prochaines analyses.

4. Conclusion

En matière d'analyse de données, il convient de respecter les différentes étapes de préparation pour faciliter la réalisation des travaux. En effet, l'expérience a montré que les phases amont sont souvent plus consommatrices de temps mais conditionnent la réussite des travaux d'analyse de données. À titre indicatif, sur une mission classique de l'Audit Informatique où des travaux d'analyse de données sont menés, la charge de travail se répartit sur les différentes phases du processus de la façon suivante :

Phases du processus	% de la charge de travail totale
Phase 1 : Prise de connaissance	10 %
Phase 2 : Définition des scénarios d'analyse	10 %
Phase 3 : Récupération des données	15 %
Phase 4 : Préparation des données	15 %
Phase 5 : Réalisation des tests	30 %
Phase 6 : Interprétation et présentation des résultats	20 %

Par conséquent, il faut se fixer des objectifs clairs et stables tout au long de la mission, également maintenir une bonne communication en interne mais aussi avec les audités, travailler par étapes pour la réalisation des tests en s'appliquant à contrôler et documenter les résultats obtenus.

En revanche, il faut veiller à ne pas se lancer dans un défi technique. Il ne s'agit pas de réécrire l'application, ni d'en faire la recette fonctionnelle. L'équipe doit être également vigilante par rapport aux délais car certains tests peuvent s'avérer plus longs que prévu d'où la nécessité de prioriser. Attention à l'interprétation hâtive, situation dangereuse si l'auditeur n'a pas toute la connaissance fonctionnelle. Enfin, l'absence de contrôles et de documentation induit le risque de fonder ses conclusions sur des résultas erronés ou d'être dans l'impossibilité de justifier la démarche d'obtention du résultat.

L'intégration d'outils d'analyse de données dans les travaux d'audit a donc nécessité un investissement important appuyé par la direction de l'audit informatique.

5. Exemples d'utilisation

a) Vérification de la mise en œuvre d'une procédure de saisie par les utilisateurs

Objectif de l'analyse : Vérification de la mise en œuvre d'une procédure de saisie par les utilisateurs qui prévoyait la saisie d'un évènement pour marquer chacune des deux étapes constitutives d'une opération manuelle. Le respect de cette procédure permettait d'offrir une visibilité sur l'état d'avancement de l'opération considérée.

Population observée : Ensemble des journaux de transactions de l'application sur une semaine sous la forme de fichiers d'évènements au format texte.

Travaux réalisés :

• Concaténation des fichiers évènements.

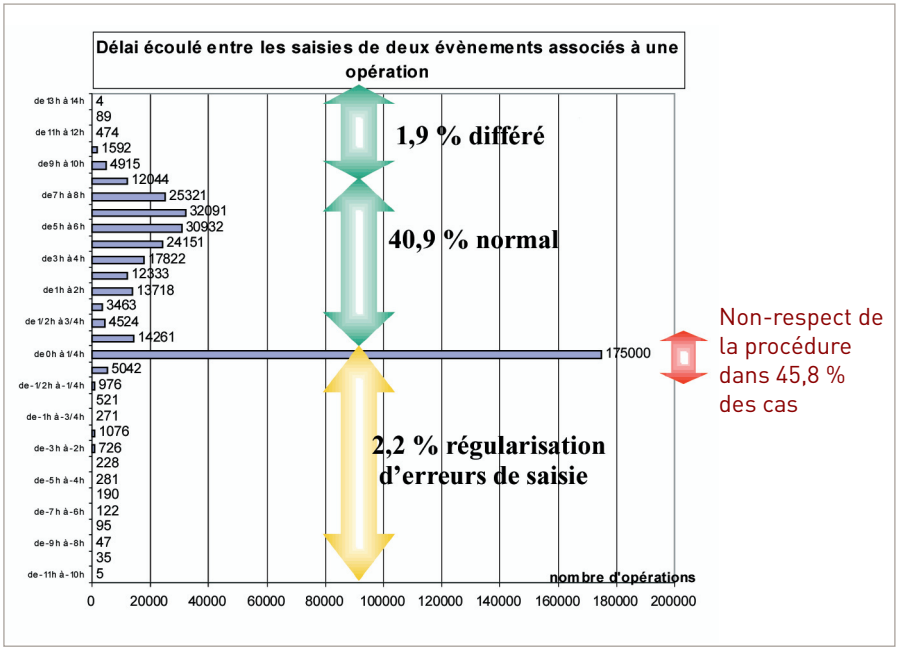
• Remise en forme globale pour importation du fichier dans l'outil d'analyse, ce qui a constitué l'étape la plus longue et la plus fastidieuse.

• Importation du fichier obtenu.

• Pour chaque opération, calcul du délai écoulé entre les deux évènements.

Résultats :

L'analyse des délais écoulés entre les deux évènements a permis de mettre en évidence un non-respect de la procédure sur le terrain et de quantifier l'étendue du dysfonctionnement. En effet, certains utilisateurs effectuaient les deux saisies en simultanée plutôt qu'à chaque fin d'étape constitutive, faussant la vue restituée par le système. À la suite de cet audit, un travail réalisé conjointement avec le propriétaire du processus a permis de mettre en œuvre les actions nécessaires pour un retour à une situation maîtrisée et conforme à la procédure.



b) Étude de l'efficacité d'un contrôle informatique additionnel mis en œuvre dans une application

Objectif de l'analyse : Démontrer l'efficacité du contrôle informatique.

Population observée : Extraction de l'ensemble des enregistrements rejetés dans l'application au cours d'un trimestre sous la forme d'un fichier Excel.

Travaux réalisés :

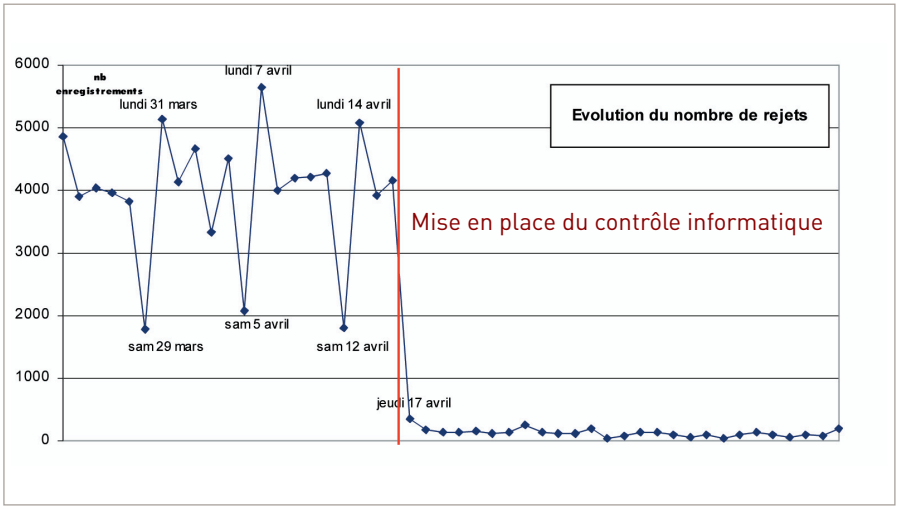
• Importation du fichier.

• Sommaire par date du nombre de rejets.

• Représentation du résultat sous forme graphique.

Résultats :

Dans ce cas, l'analyse a permis de démontrer l'efficacité du contrôle informatique puisque sa mise en place a engendré une diminution significative du nombre de rejets dans l'application. De plus, ces travaux ont montré la nécessité d'instaurer une procédure d'analyse systématique des rejets afin de détecter les dysfonctionnements éventuels notamment après la mise en production d'évolutions de l'application. À la suite de cet audit, le propriétaire du processus a donc mis en place une procédure de suivi des anomalies avec un indicateur d'alerte basé sur les tables de rejets de l'application.



c) Contrôle des habilitations par rapprochement avec les bases RH

Objectif de l'analyse : Vérification de la légitimité des habilitations accordées au niveau d'une application.

Population observée : Extraction de l'ensemble des habilitations ainsi que des droits accordés au niveau de l'application sous la forme d'un fichier Excel.

Travaux réalisés :

- Importation du fichier des habilitations dans l'outil d'analyse.
- Extraction de la liste des utilisateurs concernés.
- Extraction à partir des bases RH du statut des utilisateurs impactés.
- Importation du fichier des utilisateurs dans l'outil d'analyse.
- Rapprochement des statuts et des habilitations pour chaque utilisateur.

Résultats :

L'analyse a permis d'identifier des cas où la procédure de retrait des habilitations n'avait pas été respectée au niveau local. L'audit a abouti au lancement d'un plan de sensibilisation auprès des managers locaux afin qu'ils s'assurent du respect de la procédure.

d) Qualité de la recette fonctionnelle d'une application

Objectif de l'analyse : Vérification de la qualité de la recette fonctionnelle précédant la mise en production de la nouvelle version d'une application par examen des anomalies fonctionnelles et des opérations de maintenance à chaud.

Population observée : Extraction de l'ensemble des anomalies fonctionnelles et des opérations de maintenance à chaud sur 10 mois sous la forme de fichiers Excel.

Travaux réalisés :

- Importation des fichiers dans l'outil d'analyse.
- Cumul par date du nombre d'anomalies fonctionnelles.
- Cumul par date du nombre d'opérations de maintenance à chaud.
- Représentation des résultats sous forme graphique.

Résultats :

L'analyse a permis d'étayer les éléments recueillis au cours des entretiens d'audit. En effet, les résultats montrent que la mise en production de la nouvelle version a engendré une croissance importante du nombre d'anomalies fonctionnelles et par conséquent d'opérations de maintenance à chaud. L'analyse a donc permis de souligner l'insuffisance de la recette fonctionnelle réalisée avant la mise en production.

e) Mise à jour de la modélisation d'une activité

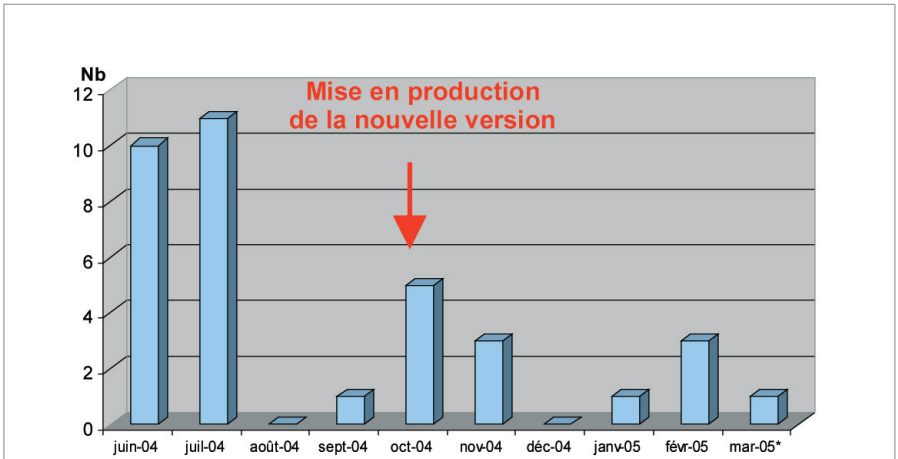
Objectif de l'analyse : Vérification du respect d'une procédure prévoyant la mise à jour annuelle de la modélisation d'une activité dans une application. Le respect de cette procédure permettait à l'application de restituer une image fiable de l'activité globale.

Population observée : Extraction de l'ensemble des modélisations créées dans l'outil au niveau national dans un fichier au format csv. Chaque modélisation était rattachée à une entité et possédait une date de création.

Travaux réalisés :

- Importation du fichier des modélisations dans l'outil d'analyse.

> Évolution du nombre d'anomalies fonctionnelles



> Évolution du nombre d'opérations de maintenance à chaud

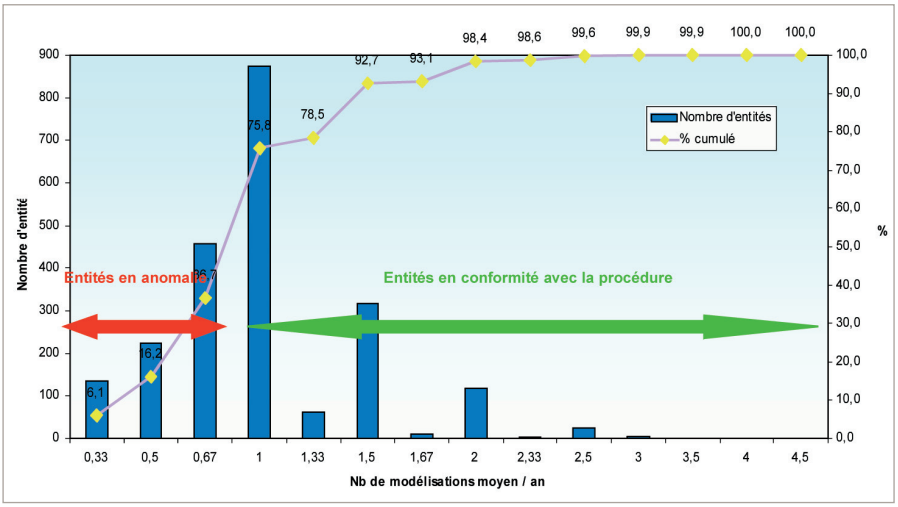
• Pour chaque entité, calcul du nombre moyen de modélisations réalisées chaque année (taux de mise à jour).

• Cumul par taux de mise à jour du nombre d'entités.

• Représentation du résultat sous forme graphique.

Résultats :

L'analyse a permis de constater que des entités ne mettaient à jour leur modélisation chaque année. Ce groupe d'entités non négligeable ne respectait pas la procédure. Une série d'actions a été déclenchée par la suite pour s'assurer de cette mise à jour annuelle dans l'ensemble des établissements.



6. Webographie

www.afai.fr :
AFAI (Association Française de l'Audit et du Conseil Informatiques)

www.isaca.com :
ISACA (Information Systems Audit and Control Association)

www.ifaci.com :
IFACI (Institut Français de l'Audit et du Contrôle Interne)

http://www.theiia.org/itaudit :
site dédié à l'audit des systèmes d'information

www.auditnet.org :
Audit Net, site dédié au partage de connaissance des auditeurs

www.caseware-idea.com :
Distributeur du logiciel IDEA

www.acl.com :
Distributeur du logiciel ACL

www.cncc.fr :
CNCC Compagnie Nationale des Commissaires aux Comptes ■

Que vaut mon système d'information ? Tout et rien. Mais peut-être la question est-elle mal posée. En effet, un système d'information est à la fois un outil de gestion du quotidien, un levier d'évolution de l'entreprise et un instrument de sa stratégie. Alors peut-être faut-il pour répondre à la question de la valeur du SI s'interroger d'abord sur sa contribution à la valeur de l'entreprise. Tâche apparemment très complexe, mais en réalité assez simple à réaliser. Pourvu que tous les protagonistes (directions générales, métiers, DSI,...) s'y retrouvent et partagent un langage commun. C'est le seul et unique but de cet ouvrage, qui n'a pas la prétention d'être une méthode, mais simplement un guide de bonne pratique, et une aide à construire une démarche. Ce n'est qu'un jalon, et la route est longue.